



浙江大学
ZHEJIANG UNIVERSITY

工作总结



引入

- 近半年的工作主要围绕以下两个Idea进行：
- 稀疏降采样
- 适应性四叉树卷积



稀疏降采样

- 稀疏降采样是用于代替传统池化的降采样方式，主要针对目标区域占比小、分布集中的情况进行改进。
- 它首先对特征图进行分块，并进行稀疏化处理，抑制那些含较少有效信息的块。
- 然后利用可训练的评估器，从中选取特定比例的区块作为降采样的输出。



稀疏降采样

- 要得到减半尺寸的输出，稀疏降采样首先将输入分为多个同样尺寸的patch，然后它们的四分之一会根据评估器得到的对应特征分选取出来并分配到合适的位置，拼接得到最终的输出
- 如将输入I 先分为N个同尺寸patches P_1, P_2, \dots, P_N ，输出的形式就将如下：
$$\begin{bmatrix} P_{i1} & \dots & P_{ik} \\ \vdots & \ddots & \vdots \\ P_{ij} & \dots & P_{i\frac{N}{4}} \end{bmatrix}$$
- $i1, i2, \dots, i\frac{N}{4}$ 是待定的



稀疏降采样

- 上面所提到的评估器 $E(\cdot)$ 为一个网络，它将所有 N 个patch作为输入，对于每个patch P_k 它给出 $\frac{N}{4}$ 个特征分:

$$S_{11}^k, S_{12}^k, \dots, S_{1\frac{\sqrt{N}}{2}}^k, \dots, S_{\frac{\sqrt{N}}{2}1}^k, \dots, S_{\frac{\sqrt{N}}{2}\frac{\sqrt{N}}{2}}^k$$

- 对应矩阵的 $\frac{N}{4}$ 个位置
- 最终的输出可以表示如下:

$$\sum_{i,j} I_i \left(\sum_k \text{select}(S_{ij}^k - \max_l \{S_{ij}^l\}) P_k \right) I_j^T$$
$$i, j \in \left\{1, \dots, \frac{\sqrt{N}}{2}\right\}, k, l \in \{1, \dots, \frac{N}{4}\}$$



稀疏降采样

- I_k 的形式如下:

$$I_k = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ I \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

第k块为单位阵，其余为零

- $\text{select}(\cdot)$ 为分段函数，在定义后为可微的



稀疏降采样

$$\text{select}(x) = \begin{cases} 1, & x = 0 \\ 0, & x < 0 \end{cases}, \text{select}'(x) = \begin{cases} 0, & x = 0 \\ 0.1, & x < 0 \end{cases}$$

稀疏化

- 稀疏化由额外增加的损失函数完成：

$$e^{-\sum_i \text{rectify}(\|P_i\|)}$$

- $\|\cdot\|$ 为矩阵的1-范数， $\text{rectify}(\cdot)$ 为含有一个阈值 ϵ 的分段函数：

$$\text{rectify}(x) = \begin{cases} -\epsilon^{-1}x + 1, & 0 \leq x < \epsilon \\ -0.01x + 0.01\epsilon, & \epsilon \leq x \end{cases}$$



稀疏降采样

- 在LIDC数据集上的对比实验表明新的降采样方式可以提高约2%原有网络的分类精度.

网络	分类准确率	AUC
CNN	85.61	89.41
CNN+4SD	87.71	91.44
CNN+6SD	87.45	91.29

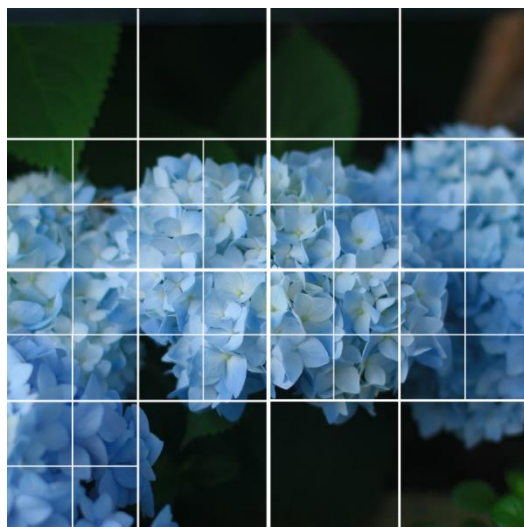
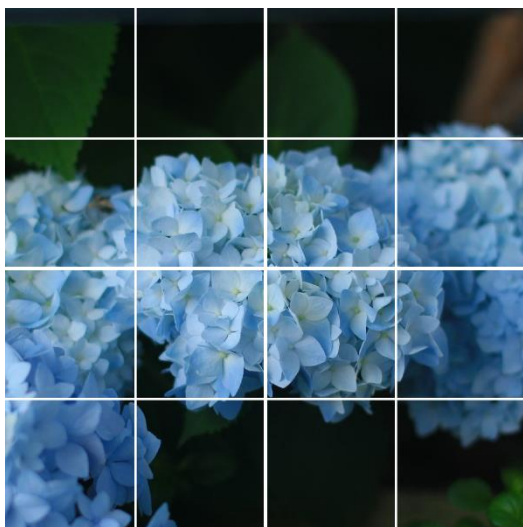


适应性四叉树卷积

- 适应性四叉树卷积是对传统卷积方式的改进。
- 传统卷积对图像的每一部分都进行了同等规模的计算，有很相当部分计算用于背景等信息较为稀疏的区域中。
- 它基于四叉树的思想，适应性地对图像不同部分进行不同程度的卷积操作。



适应性四叉树卷积



适应性四叉树卷积

- 通常的卷积以如下方式进行，将形如 (C_{in}, H, W) 的图像输出为 $(C_{out}, H_{out}, W_{out})$ 的特征图：

$$\text{out}(C_{outj}) = \text{bias}(C_{outj}) + \sum_{k=0}^{C_{in}-1} \text{weight}(C_{outj}, k) \star \text{input}(k)$$

- 适应性四叉树卷积对先按照梯度信息对图像进行N轮四叉树分割，每一轮分割都在上轮超过阈值的部分上进行，对变化程度大的部分进行精细分割。



适应性四叉树卷积

- 每轮都得到原图像的一个子集，共N个即

S_1, S_2, \dots, S_N .

- 对应的适应性四叉树卷积在每个通道上都有N个卷积核，上述的N个集合规定了这N个卷积核的作用范围。
- 以如下方式进行：

$$\text{out}(C_{out_j}) = \text{bias}(C_{out_j}) + \sum_{k=0}^{C_{in}-1} \text{weight}^N(C_{out_j}, k) \dots \text{weight}^1(C_{out_j}, k) \star S_1(k) \dots S_N(k)$$



适应性四叉树卷积

- 每轮卷积迭代都仅在图像的一部分上进行，且区域在不断缩小。
- 现有两个问题：
- 对图像的评估（用于阈值选取）开始由网络完成，但这极大增加了计算量，所以改为更简单直接的方式，造成了欠缺说服力的问题。



适应性四叉树卷积

- 新的卷积方式无法使用现有深度学习框架完全实现，造成了目前的实验只能进行单轮分割，在LIDC数据集上的实验结果仅有80%左右的分类精度。
- 解决方法只能是从底层来实现新的卷积方式。



谢谢！

